GRUNDLAGEN DER BAYESSCHEN METAANALYSE

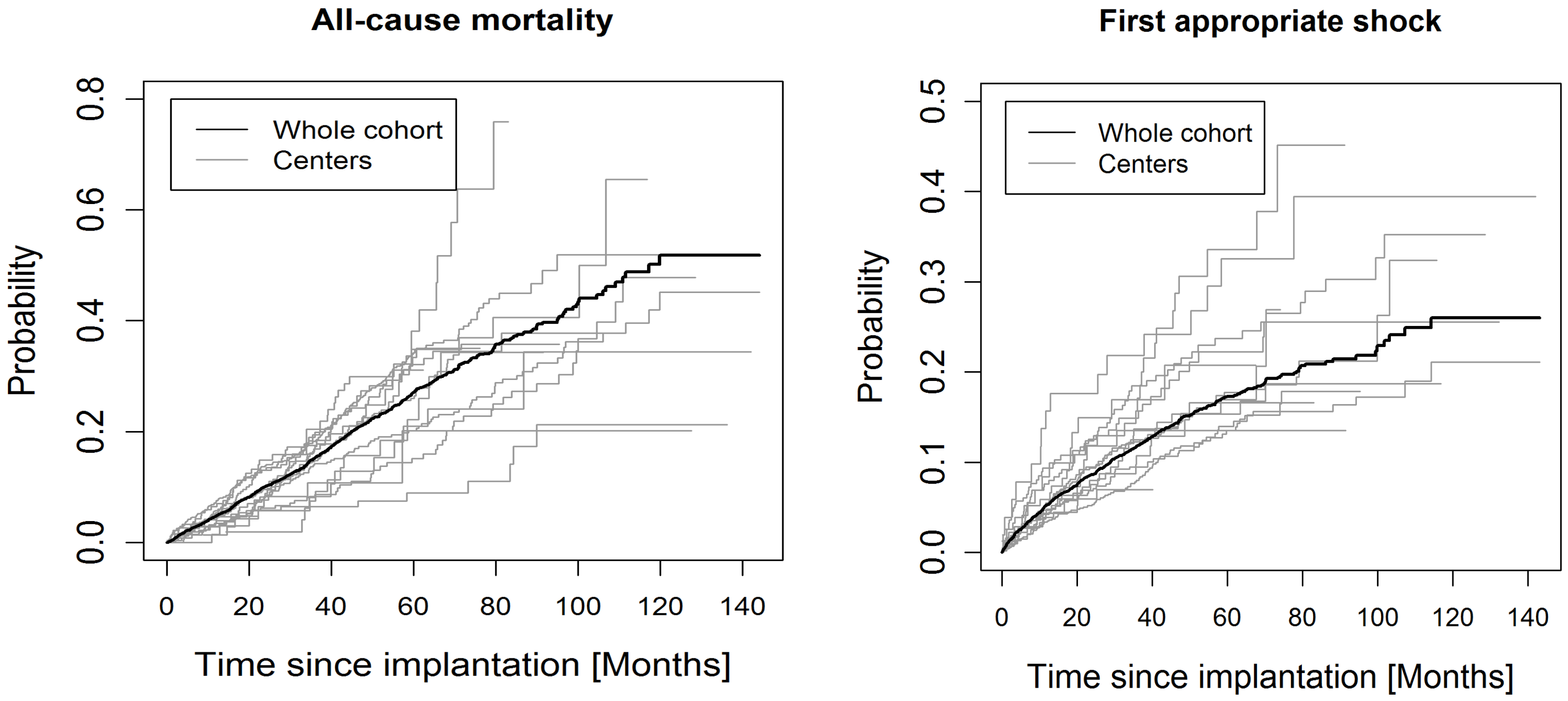
GMDS 2023: Workshop ATF und Präsidiumskommission "Methodenaspekte in der Arbeit des IQWiG und IQTIG"

Tim Friede

Institut für Medizinische Statistik

Universitätsmedizin Göttingen

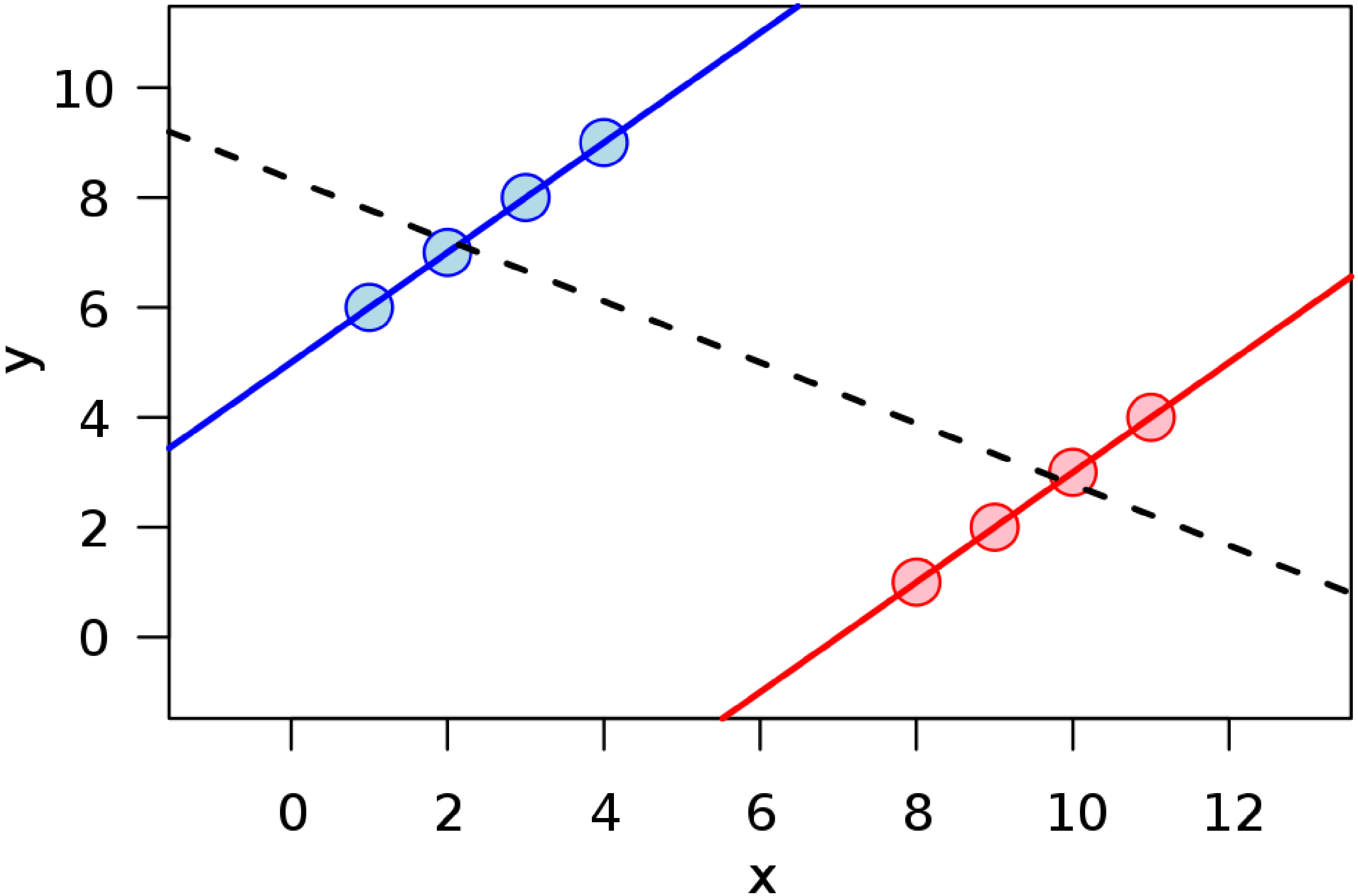
HETEROGENEITY IN EVENT RATES: EU-CERT-ICD REGISTRY



Sticherling et al (2018) Europace

SIMPSON'S PARADOX

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG



http://en.wikipedia.org/wiki/Simpson%27s_paradox

META-ANALYSIS

- **Data:** treatment effect estimate y_i and standard error σ_i for study i
- **General common-effect (or fixed-effect) model**
  - Assumption: The true (unknown) treatment effects $\theta_1, \ldots, \theta_k$ in studies 1 to k are the same (i.e. $\theta_1 = \ldots = \theta_k = \mu$).
  - Treatment effect estimate (with weights w_i): $\hat{\mu} = \frac{\sum_{i=1}^{k} w_i y_i}{\sum_{i=1}^{k} w_i}$
  - Inverse-variance weighted method: $w_i = 1/\sigma_i^2$ (with variance σ_i^2)
  - Confidence interval $\hat{\mu} \pm z_{1-\alpha/2}\sqrt{1/\left(\sum w_i\right)}$
- Some **specific methods** for combining odds ratios (e.g. Mantel-Haenszel, Peto)

EU-CERT-ICD REGISTRY: HETEROGENEITY IN GENDER DIFFERENCES?

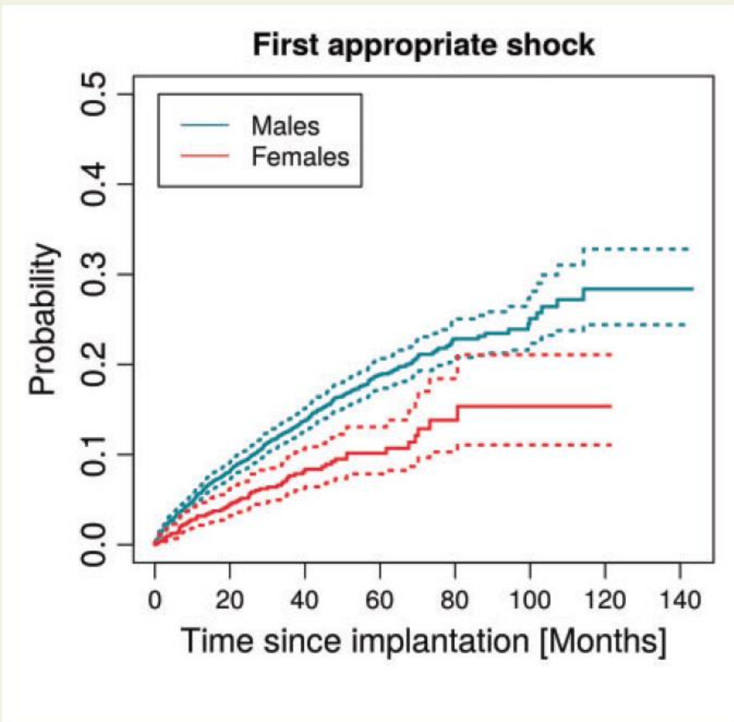


Figure 5 Cumulative incidences of first appropriate shocks by gender (with 95% CIs).

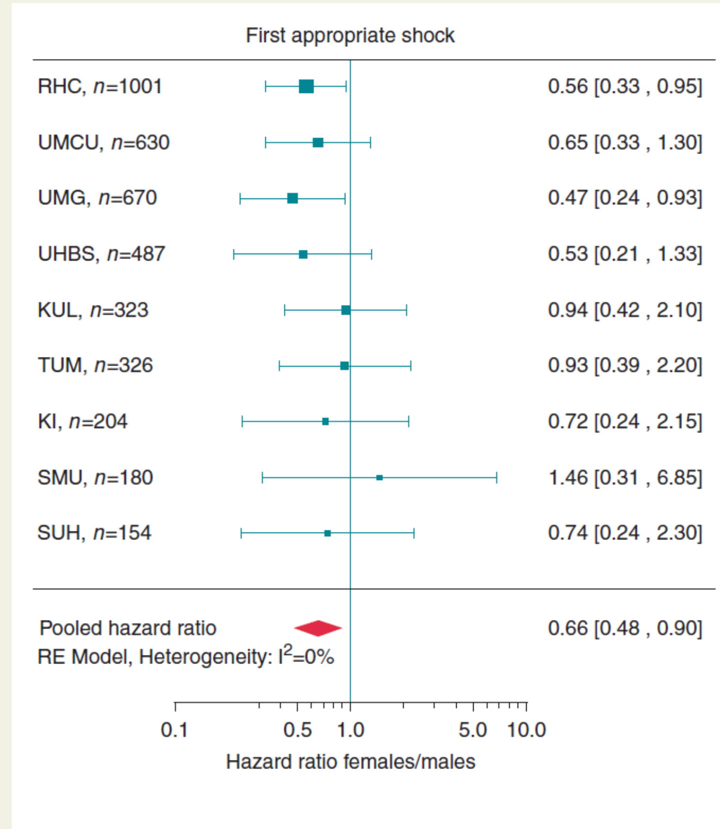


Figure 6 Forest plot of estimated centre-specific hazard ratios for gender regarding the first appropriate shock together with their 95% CIs and the pooled hazard ratio with a modified Knapp-Hartung 95% CI. (Note that only centres with at least 10 female patients and at least 1 observed first appropriate shock both among males and females were included in this analysis.).

Sticherling et al (2018) Europace

EU-CERT-ICD SYSTEMATIC REVIEW: HETEROGENEITY IN GENDER EFFECTS

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

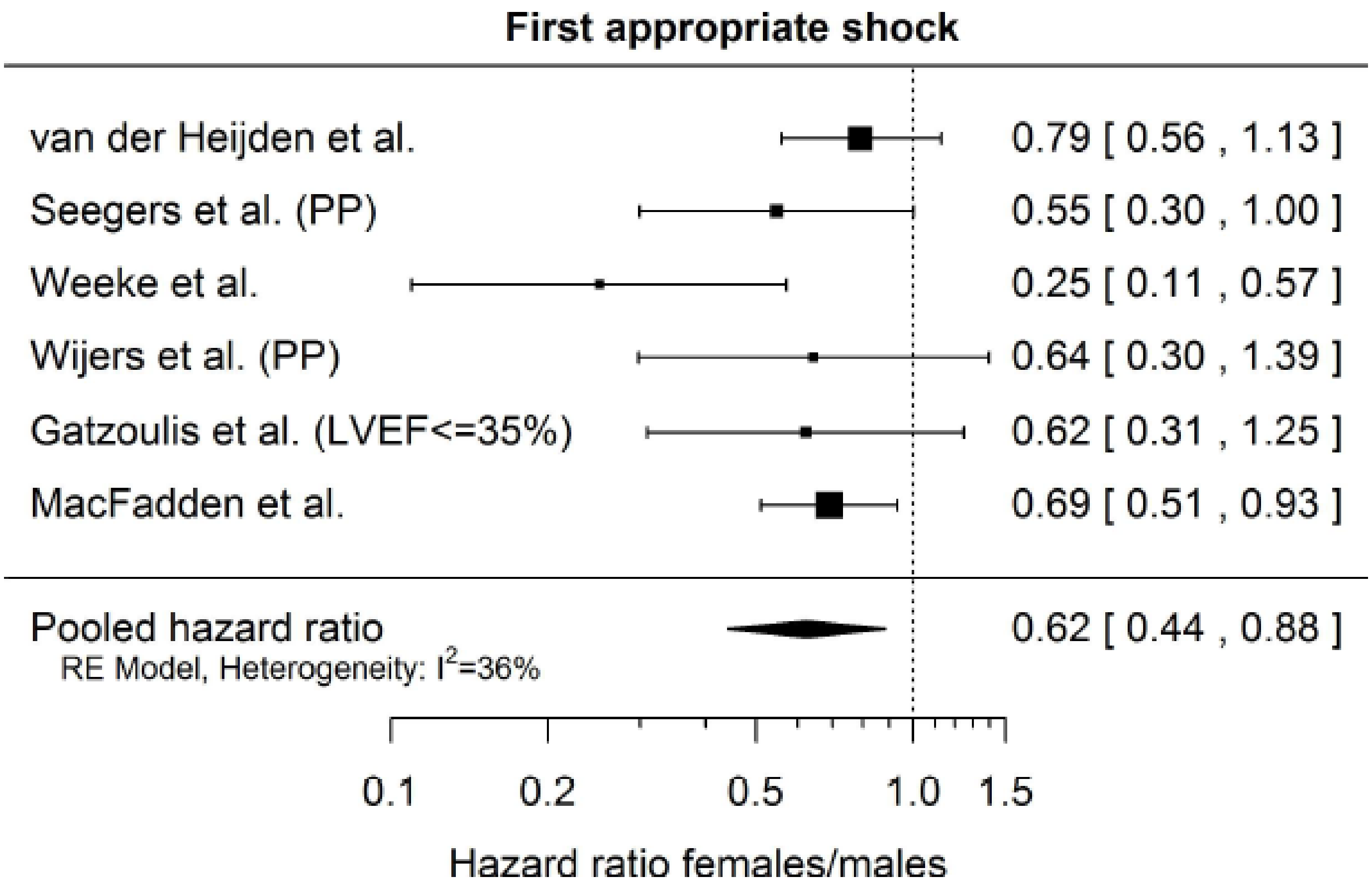


Fig 3. Extracted hazard ratios for female gender regarding risk of all-cause mortality with 95% confidence intervals as reported in the respective publications. 'PP' indicates that the results were re-analyzed for primary prevention patients only. The pooled estimate is reported with a Knapp-Hartung adjusted 95% confidence interval. The dotted vertical line denotes a hazard ratio of 1, which corresponds to no difference in the risk between males and females.

Conen et al (2016) PLoS ONE

NORMAL-NORMAL HIERARCHICAL MODEL (NNHM)

- **Common effect (or fixed effect) model**
  - assumes no between-study heterogeneity (i.e. $\theta_1 = \ldots = \theta_k$)
  - confidence intervals too narrow if heterogeneity present
- **NNHM for random effects meta-analysis**
  - Study-specific effect sizes $\theta_1, \ldots, \theta_k$ from a normal distribution with mean μ and variance τ^2, i.e. $\theta_i | \mu, \tau \sim N(\mu, \tau^2)$
  - Therefore, $y_i | \mu, \tau \sim N(\mu, \sigma_i^2 + \tau^2)$
  - Hence, the weights become $w_i = 1/(\tau^2 + \sigma_i^2)$
  - Formulae for the overall treatment effect and its standard error the same as for the common effect model, but with different weights (see above)

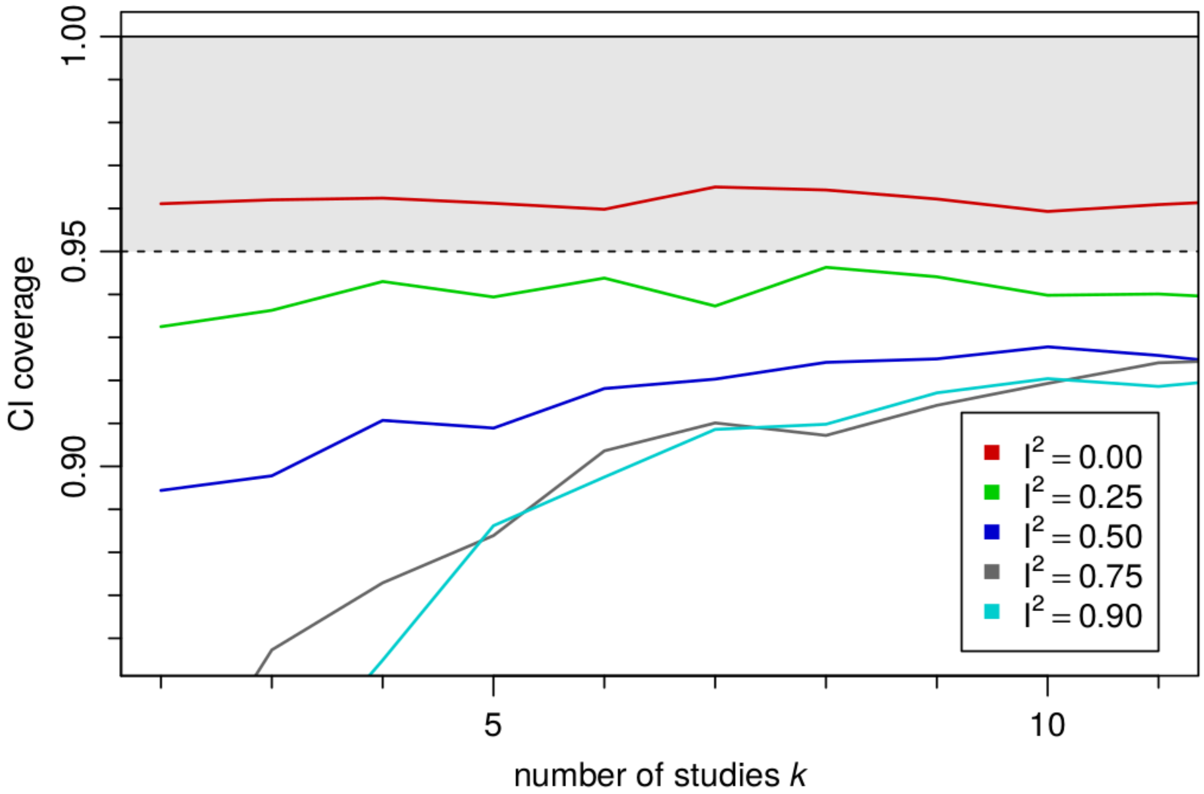
BETWEEN-STUDY HETEROGENEITY

- **Between-study heterogeneity and how we dealt with it**
  - Baseline / control arm (e.g. event rate): stratification by study
  - Treatment effects: random effects meta-analysis
- **Meta-analyses including only (very) few studies common**
  - Cochrane database: meta-analyses of 2-3 studies very common (Turner et al, 2012)
  - Summarizing studies of a development programme
  - …

STANDARD METHOD FAILS

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

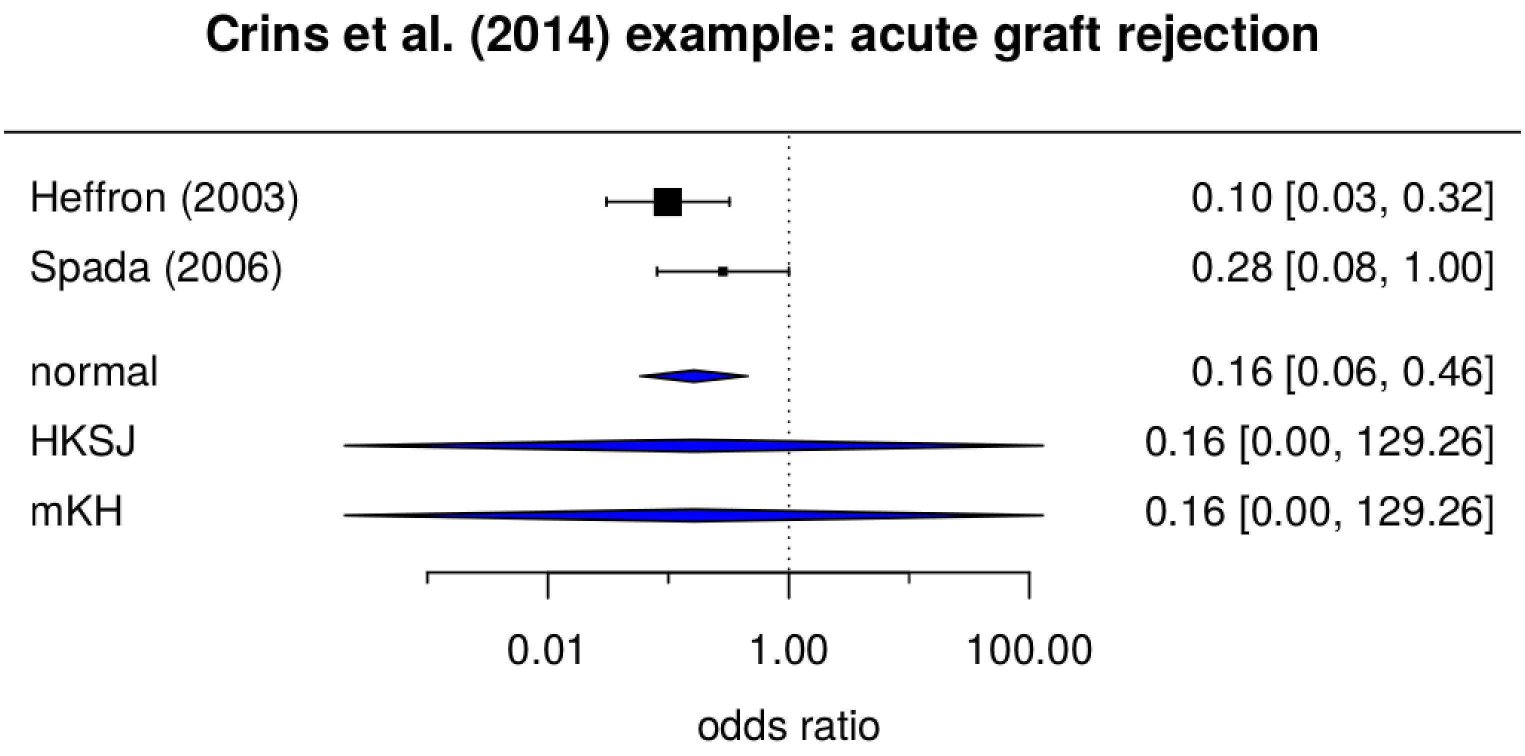
- Standard method (DerSimonian-Laird, DL)
  - Underestimates between-study heterogeneity
  - Fails to account for uncertainty in estimation of heterogeneity



IntHout et al, 2014; Röver et al, 2015

WITH VERY FEW STUDIES: KNAPP-HARTUNG METHOD DOES NOT SOLVE THE PROBLEM

- 97.5% quantile of t-distribution with 1 df = 12.7 !!!
- Example from Friede et al (2017b)



HSJK: Hartung-Knapp-Sidik-Jonkman; mHK: modified Knapp-Hartung; normal: DL

RANDOM EFFECTS META-ANALYSES WITH (VERY) FEW STUDIES

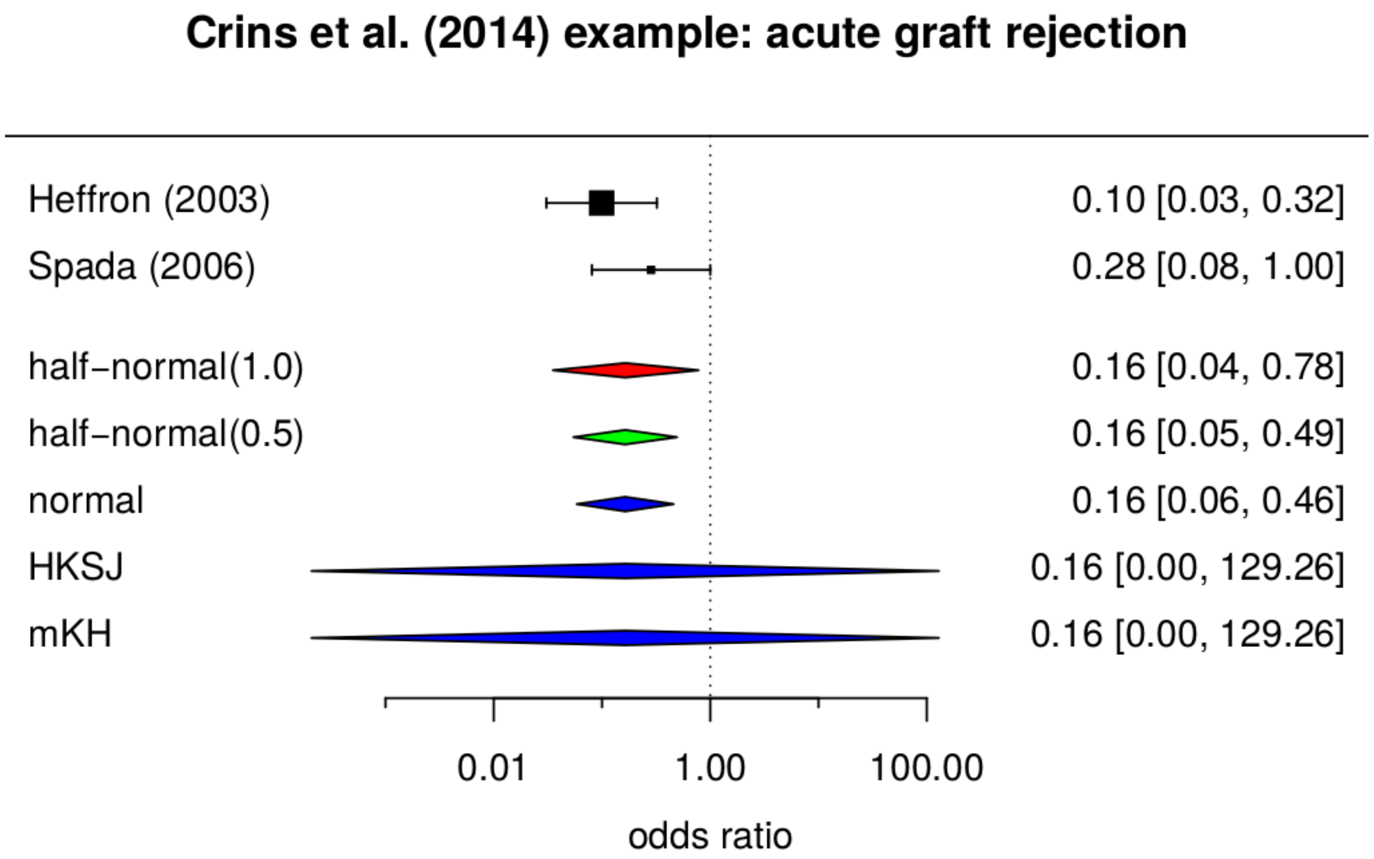
- **Standard methods** (using normal approximation)
  - confidence intervals too short; do not have the right coverage
- **Extensions based on t-distributions and rescaling of standard errors** (e.g. Knapp-Hartung method)
  - good coverage if the standard errors form different studies similar
  - in general, however, HKSJ intervals either so wide that they do not allow any conclusion, or very narrow. The latter occurs rarely, but can lead to problematically narrow confidence intervals and unfavourable coverage.
- **Bayes random-effects meta-analysis …**

BAYESIAN META-ANALYSIS

- **Idea**: Weakly informative prior on between-trial heterogeneity for meta-analysis with few studies (Spiegelhalter et al, 2004), with uninformative prior on treatment effect
  - Avoids zero estimates of between-trial heterogeneity
  - Accounts for uncertainty in the estimation of the heterogeneity
- **Easy to compute**
  - Application of DIRECT algorithm (Röver & Friede, 2017) (which is faster than MCMC sampling and does not require inspection of convergence diagnostics)
  - R package `bayesmeta` by Christian Röver (available from CRAN)

EXAMPLE REVISITED

- **Bayesian intervals** appear to be a reasonable compromise (supported by simulation studies in e.g. Friede et al, 2017a,b)



PRIORS COVERING SMALL TO LARGE HETEROGENEITY ON LOG-ODDS RATIO SCALE

Table 1. Between-trial heterogeneity for log-odds ratios: τ values representing small to very large heterogeneity, with 95% intervals for across-trial odds ratios ($\exp(\theta_j)$).

Heterogeneity		95% interval
Small:	$\tau = 0.125$	0.783–1.28
Moderate:	$\tau = 0.25$	0.613–1.63
Substantial:	$\tau = 0.5$	0.325–2.66
Large:	$\tau = 1$	0.141–7.10
Very large:	$\tau = 2$	0.020–50.4

Table 2. Between-trial heterogeneity for log-odds ratios: three priors covering small to large heterogeneity.

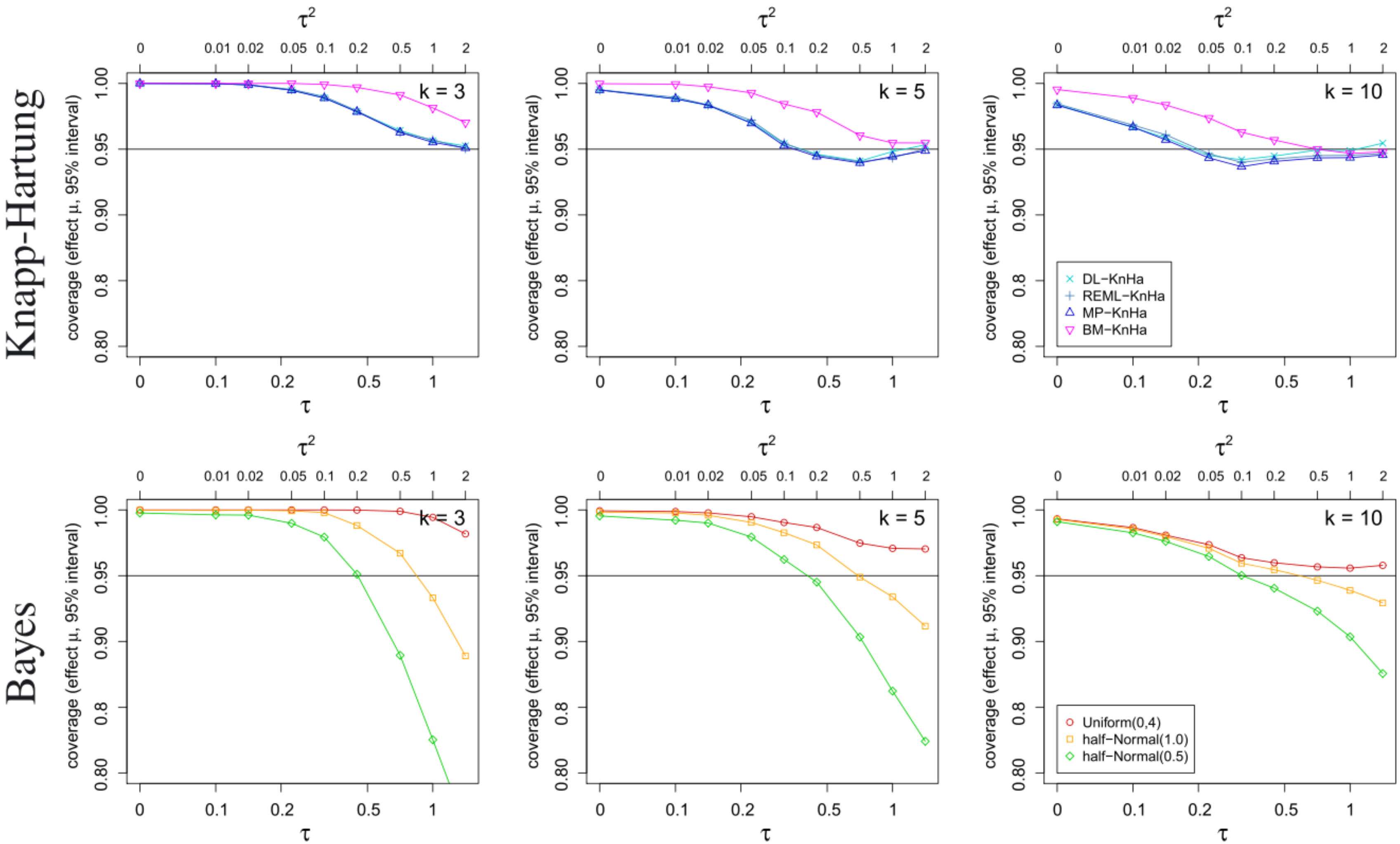
Prior distribution	Median	95% interval
Half normal (scale = 0.5)	0.337	(0.016, 1.12)
Half normal (scale = 1.0)	0.674	(0.031, 2.24)
Uniform (0, 4)	2.0	(0.1, 3.9)

Friede et al. (2017) RSM

COVERAGE PROBABILITY

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

- Coverage for confidence / credibility intervals of overall effect

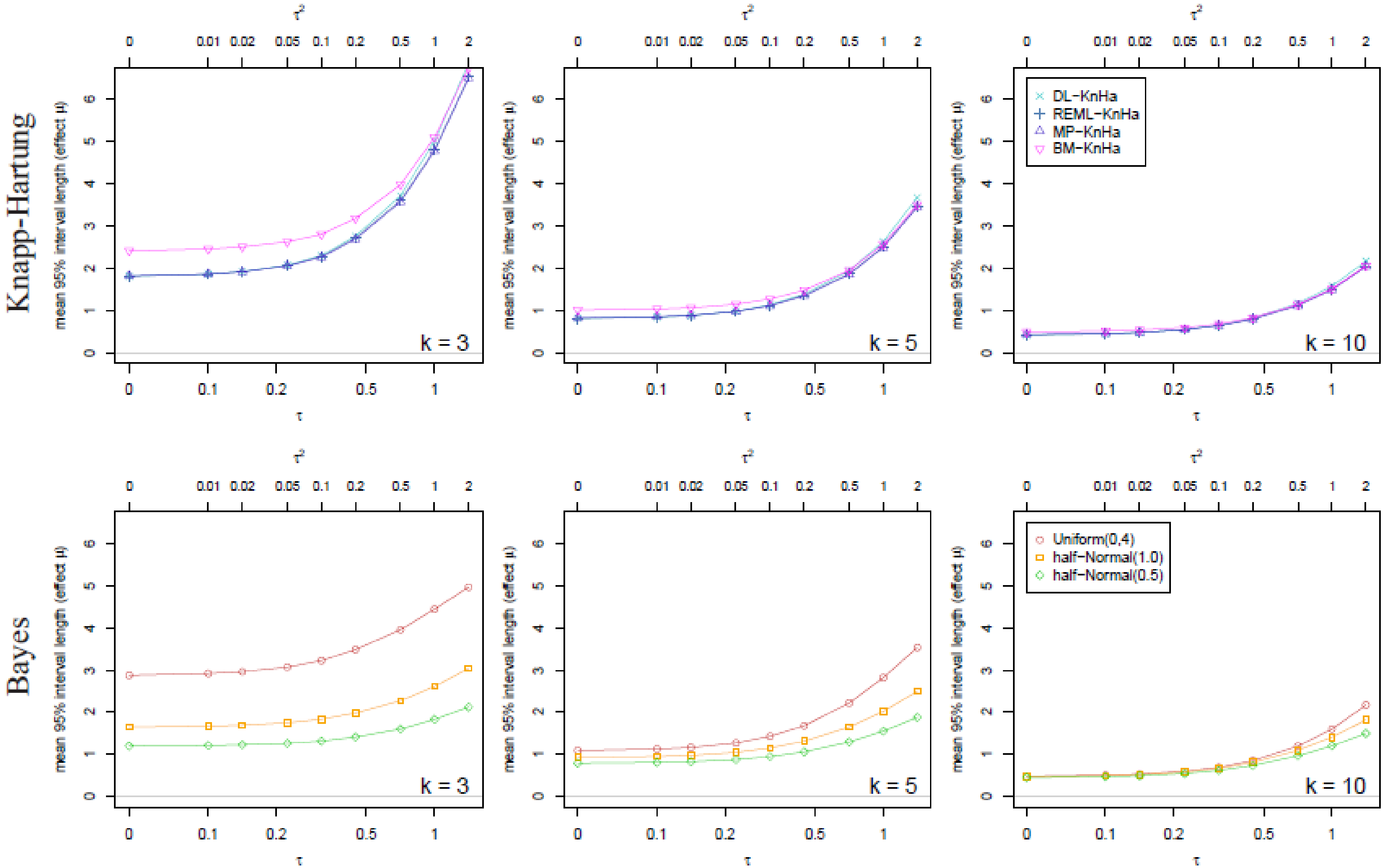


Friede et al. (2017) RSM

BETWEEN-TRIAL HETEROGENEITY

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

▷ Mean length of confidence / credibility intervals



Friede et al. (2017) RSM

“WHERE DOES THE PRIOR COME FROM?”

- Theoretical arguments, simulations, data

Received: 16 July 2020 | Revised: 13 January 2021 | Accepted: 16 January 2021
DOI: 10.1002/jrsm.1475
RESEARCH ARTICLE
Research Synthesis Methods WILEY

On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis

Christian Röver[1] | Ralf Bender[2] | Sofia Dias[3] | Christopher H. Schmid[4] | Heinz Schmidli[5] | Sibylle Sturtz[2] | Sebastian Weber[6] | Tim Friede[1]

Received: 25 February 2022 | Revised: 16 November 2022 | Accepted: 18 March 2023
DOI: 10.1002/sim.9731
RESEARCH ARTICLE
Statistics in Medicine WILEY

Summarizing empirical information on between-study heterogeneity for Bayesian random-effects meta-analysis

Christian Röver[1] | Sibylle Sturtz[2] | Jona Lilienthal[2] | Ralf Bender[2] | Tim Friede[1]

SOME REFERENCES

- Friede T, Röver C, Wandel S, Neuenschwander B (2017a) Meta-analysis of few small studies in orphan diseases. Research Synthesis Methods 8: 79–91.
- Friede T, Röver C, Wandel S, Neuenschwander B (2017b) Meta-analysis of two studies in the presence of heterogeneity with applications in rare diseases. Biom J 59: 658 – 671.
- Röver C, Friede T (2017) Discrete approximation of a mixture distribution via restricted divergence. Journal of Computational and Graphical Statistics 26: 217-222.
- Röver C, Knapp G, Friede T (2015) Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. BMC Medical Research Methodology 15: 99.